

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Інститут телекомунікаційних систем
Кафедра Інформаційно-телекомунікаційних мереж**

До захисту допущено:

Завідувач кафедри

_____ Лариса ГЛОБА

«__» _____ 2020 р.

Дипломна робота
на здобуття ступеня бакалавра
за освітньо-професійною програмою «Інформаційно-комунікаційні
технології»
спеціальності 172 «Телекомунікації та радіотехніка»
на тему: «5G, швидкісна передача даних в мережі»

Виконав:

студент IV курсу, групи ПІ-62

Драбкін Денис Дмитрович _____

Керівник:

доцент кафедри ІТМ ІТС,

Алексєєв Микола Олександрович, к.т.н., с.н.с _____

Рецензент:

Посада, науковий ступінь, вчене звання,

Прізвище, ім'я, по батькові _____

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент _____

Київ – 2020 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Інститут телекомунікаційних систем
Кафедра Інформаційно-телекомунікаційних мереж

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 172 «Телекомунікації та радіотехніка»

Освітньо-професійна програма «Інформаційно-комунікаційні технології»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Лариса ГЛОБА

«__» _____ 2020 р.

ЗАВДАННЯ

на дипломну роботу студенту

Драбкіну Денису Дмитровичу

1. Тема роботи «5G, швидкісна передача даних в мережі», керівник роботи доцент кафедри інформаційно-телекомунікаційних мереж ІТС Алексєєв Микола Олександрович, к.т.н., с.н.с., затверджені наказом по університету від «30» березня 2020 р. № 924-с.

2. Термін подання студентом роботи 10 червня 2020 р.

3. Вихідні дані до роботи:

1. Технологія CDN, як можливість поширення контенту між пристроями;
2. Edge computing – технологія периферійного обчислення;

4. Зміст роботи:

1. Провести аналітичний огляд технології CDN;
2. Провести огляд технології периферійного обчислення Edge Computing у мобільних мережах, зокрема у мережі 5G;

3. Огляд можливостей мережі 5G для побудови CDN для поширення контенту між мобільними пристроями;
5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо)
6. Дата видачі завдання 20 листопада 2019 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	Дослідження основних задач, котрі вирішує технологія CDN;	20.11.19 – 09.01.20	виконано
2	Аналіз принципу роботи CDN	02.01.20 – 12.03.20	виконано
3	Огляд технології Edge computing	13.03.20 – 23.04.20	виконано
4	Дослідження застосування технологій периферійних обчислень у мобільних мережах	24.04.20 – 10.05.20	виконано
5	Огляд можливостей мережі 5G для побудови CDN для поширення контенту між мобільними пристроями	11.05.20 – 26.05.20	виконано
6	Підготовка тексту диплому	27.05.20 - 08.06.20	виконано

Студент

Денис ДРАБКІН

Керівник

Микола АЛЕКСЄЄВ

ЗМІСТ

ВСТУП	9
РОЗДІЛ 1. АНАЛІТИЧНИЙ ОГЛЯД ТЕХНОЛОГІЇ CDN ДЛЯ ВИКОРИСТАННЯ В МОБІЛЬНИХ МЕРЕЖАХ	Ошибка! Закладка не определена.
1.1. Основні поняття технології CDN..	Ошибка! Закладка не определена.
1.1.1 Основні терміни.....	14
1.1.2 Призначення та сфера використання CDN.....	14
1.2. Алгоритми розподілу запитів.....	15
1.3. Принцип роботи CDN.....	Ошибка! Закладка не определена.
1.4. Аналіз постачальників мереж доставки контенту і готових рішень для їх створення.....	24
Висновки.....	27
РОЗДІЛ 2. ОГЛЯД ТЕХНОЛОГІЙ ПЕРЕФЕРІЙНОГО ОБЧИСЛЕННЯ. EDGE COMPUTING	28
2.1. Призначення технологій периферійних обчислень	28
2.2. Застосування технологій периферійних обчислень у мобільних мережах	29
2.2.1 Застосування у мережах 2G.....	31
2.2.2 Застосування у мережах 3G.....	32
2.2.3 Застосування у мережах 4G/LTE/4G+/LTE-A.....	34
2.2.4 Застосування у мережах 5G.....	35
2.3. Огляд можливостей мережі 5G для побудови CDN для поширення контенту між мобільними пристроями	40
Висновки	42
ЗАГАЛЬНІ ВИСНОВКИ ПО РОБОТІ	43
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	44

РЕФЕРАТ

Робота містить 45 сторінок, 8 рисунків та 1 таблиця. Було використано 5 джерел.

Мета роботи: підвищення ефективності обробки великих обсягів обчислюваних даних та покращення швидкості доставки контенту за рахунок технології граничного обчислення.

В ході виконання даної роботи було проаналізовано технологію доставки контенту до користувача та проблему навантаження ЦОД при великому обсязі даних та їх обробки. Розглянуто основні поняття та принципи доставки контенту(CDN). На базі проведеного аналізу запропоновано вдосконалений метод обробки даних за рахунок збільшення обсягу інформації в інтернеті, кількість якої стрімко зростає.

Виконаний аналіз можливостей мобільних мереж для використання технології периферійного обчислення.

Ключові слова: мережа доставки контенту, периферійні обчислення, граничні обчислення, мобільні мережі, 5G, edge computing.

ABSTRACT

The work consists of 45 pages, 1 table and 8 figures. Used 5 references.

Goal of diploma: to increase the efficiency of processing large amounts of calculated data and improve the speed of content delivery through the technology of boundary computing.

In the course of this work, the technology of content delivery to the user and the problem of data center load with a large amount of data and their processing were analyzed. The basic concepts and principles of content delivery (CDN) are considered. Based on the analysis, an improved method of data processing is proposed by increasing the amount of information on the Internet, the number of which is growing rapidly.

The analysis of possibilities of mobile networks for use of technology of peripheral computing is executed.

Keywords: content delivery network, peripheral computing, boundary computing, mobile networks, 5G, edge computing.

ПЕРЕЛІК СКОРОЧЕНЬ

CDN	Content Delivery Network
MNO	Mobile network operator
IP	Internet Protocol
PoP	Point of Presence
DDoS	Distributed Denial of Service
MEC	Mobile edge computing
GSM	Global System for Mobile Communications
CDMA	Code Division Multiple Access
TDMA	Time Division Multiple Access
FDMA	Frequency Division Multiple Access
SMS	Short Message Service
MMS	Multimedia Messaging Service
WAP	Wireless Application Protocol
PDA	Personal Digital Assistant
URLLC	Ultra-Reliable and Low-Latency Communication
M2M	Machine-to-Machine
RAN	Radio access network
MBMS	Multimedia Broadcast Multicast Service
CA	Carrier Aggregation
IoT	Internet of things
BW	Bandwidth availability
RC	Random Choice
HD	Hard disk availability
CON	Connection availability
HBW	Hight Bandwidth availability
MBW	Midle Bandwidth availability

LBW	Low Bandwidth availability
HHD	Hight Hard disk availability
MHD	Midle Hard disk availability
LHD	Low Hard disk availability
HCON	Hight Connection availability
MCON	Midle Connection availability
LCON	Low Connection availability

ВСТУП

Актуальність. На сьогоднішній день глобальний обсяг даних у світі складає близько 33 зеттабайт. Експерти прогнозують, що в результаті збільшення величини даних, кількості користувачів Інтернет та розвитку IoT, кількість даних складе 173 зеттабайта, які повинні бути оброблені негайно.

Такий обсяг даних є достатньо великим, тому звичайна технологія обробки Cloud Computing потребує вдосконалення або кардинально нових рішень. Методологія Edge computing(периферійні обчислення) забезпечує близькість до кінцевих пристроїв, географічний розподіл обчислюваних потужностей та зменшення навантаження на центральний ЦОД. Таким чином, граничний тип обчислень дає змогу обробки даних в межах досяжності кінцевих пристроїв не навантажуючи централізований дата-центр.

Об'єктом роботи є процес швидкісної передачі даних в мережі 5G.

Предметом роботи є технологія доставки та обробки даних в мобільних мережах.

Метою роботи є аналітичний огляд можливостей технології периферійного обчислення в мережі 5G.

Для досягнення мети дослідження було поставлено та вирішено такі

основні задачі:

1. Проведений аналітичний огляд технології CDN;
2. Проведений огляд технології периферійного обчислення Edge Computing у мобільних мережах, зокрема у мережі 5G;
3. Проведений огляд можливостей мережі 5G для побудови CDN для поширення контенту між мобільними пристроями;

Теоретичний результат дослідження:

1. Аналіз методу доставки контенту або поширення мережі мобільного контенту CDN для застосування в мобільних мережах.

2. Вдосконалений метод обробки даних для мобільних мереж, що дає можливість обчислення даних в межах кінцевих пристроїв, передаючи тільки важливу інформацію на ЦОД.

3. Дослідження можливостей мобільних мереж для використання периферійних обчислень.

РОЗДІЛ 1.

АНАЛІТИЧНИЙ ОГЛЯД ТЕХНОЛОГІЇ CDN ДЛЯ ВИКОРИСТАННЯ В МОБІЛЬНИХ МЕРЕЖАХ

Мобільна мережа доставки контенту або поширення мережі мобільного контенту (Mobile CDN) є мережею серверів - систем, комп'ютерів або пристроїв, які взаємодіють для оптимізації доставки контенту кінцевим користувачам на будь-якому типі бездротової або мобільної мережі. Як і традиційний CDN, основна мета мобільного CDN забезпечити контентом кінцевих користувачів з високою доступністю і високою продуктивністю. Крім того, мобільний CDN може бути використаний для оптимізації доставки контенту для отримання унікальних характеристик бездротових мереж і мобільних пристроїв, такі як обмежена пропускна здатність мережі або навіть розширення пристрою.

Основною ідеєю мобільних CDN є інтегрування послуги мобільної доставки, які оптимізують доставку будь-якого виду контенту, включаючи живе потокове відео за вимогою, відео і доставку інших змістовних активів. У разі відеоконтенту, ці послуги включають в себе виявлення пристроїв, рендеринг зображень та відео, транскодування та адаптація швидкості передачі. Для оптимізації доставки іншого контенту, мобільні CDN використовують такі технології, як кешування.

Один з варіантів мобільних рішень CDN, коли вузли розміщені на периферії мережі і в декількох місцях одночасно, зазвичай протягом декількох магістралей, безпосередньо пов'язаних з операторами мобільного зв'язку (MNOs). Ці вузли взаємодіють один з одним, щоб задовольнити запити кінцевих користувачів для контенту та прозоре переміщення контенту для оптимізації процесу доставки. Перевагами такої оптимізації є зниження пропускної здатності використання, підвищення продуктивності користувачів, або збільшення глобальної доступності контенту по мережі мобільного зв'язку.

Існує зростаюча потреба в підтримці продуктивності мобільних додатків за межами всього мобільного контенту доставки в браузер.

1.1. Основні поняття технології CDN

Мережа доставки (розподілу) контенту (англ. Content Delivery Network або Content Distribution Network, CDN) – географічно розподілена мережна інфраструктура, що дає змогу оптимізувати доставку і розподіл контенту кінцевим користувачам в мережі Інтернет. Використання контент-провайдерів CDN мереж сприяє збільшенню швидкості завантаження інтернет-користувачами аудіо-, відео програмного, ігрового та інших видів цифрового контенту в точках наявності мережі CDN. Основне завдання CDN – максимальне наближення контенту до кінцевого користувача.

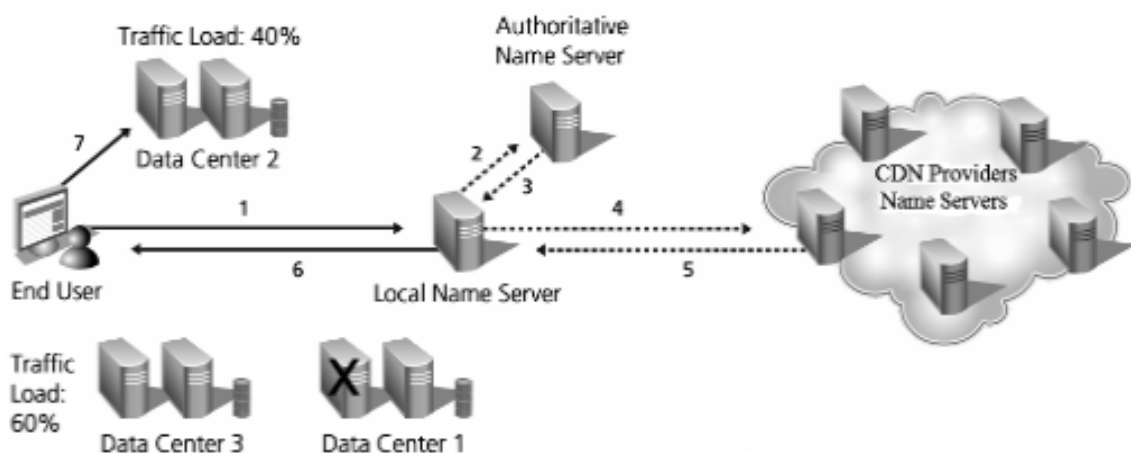


Рис 1.1 Архітектура платформи CDN

Процес отримання контенту через CDN можна описати кількома етапами (Рис 1.1):

1. Користувач запитує ресурс (веб-сайт чи відеододаток) і його браузер надсилає запит до локального DNS для отримання IP-адреси.
2. Використовуючи стандартну політику роботи служби DNS, локальний сервер запитує IP-адресу в авторитетного (англ. authoritative) DNS.
3. DNS сервер дає відповідь, включаючи запис CNAME із посиланням (ip-alias) на CDN провайдера.

4. Локальний DNS звертається до DNS сервера CDN провайдера для отримання оптимального маршруту до запитуваного контенту.

5. DNS сервери CDN провайдера віддають список адрес серверів із оптимальним маршрутом до користувача, який запитує контент.

6. Локальний DNS відправляє одну із адрес браузеру чи іншому додатку кінцевого користувача.

7. Отримавши адресу сервера, кінцевий користувач звертається за контентом до конкретного сервера, який надає його користувачеві.

Варто зазначити, що під час вибору сервера віддавання зі списку адрес, які повертають DNS сервери провайдера CDN, враховується завантаженість та доступність серверів віддавання контенту.

Основними можливостями, що надає сервіс CDN є:

1) *Потокове мовлення відео в режимі реального часу (Live Streaming)*. Можлива трансляція із сервера постачальника контенту або безпосередньо з IP-камери. Інша сфера застосування Live Streaming – трансляція прямого ефіру телеканалів через Інтернет.

2) *Потокове мовлення відео за запитом (Video on Demand Streaming)*. Передбачає перегляд відеоконтенту в будь-який момент часу, зручний кінцевому користувачу. Можливий перегляд відео з будь-якого місця без завантаження всього вмісту.

3) *Http-кешування*. Http-контент у міру потреби кешується на серверах роздавання і віддається з оптимального сервера за запитом користувача.

4) *Кодування відео в режимі реального часу*. Конвертація та перекодування одного формату файлів у інші формати чи контейнери.

5) *Транскодування відеопотоків*. Перетворення типу компресії, зміна бітрейту чи розміру відео.

Також, окрім істотного зменшення навантаження на сервери додатків, CDN мережа виконує й інші функції.

Всі сучасні провайдери розміщують копії своїх додатків на усіх

доступних для них кешуючих серверах, які перенаправляють запити користувачів до найближчих з них. Результат – значно менша затримка у доставленні даних кінцевому користувачеві та якісніше надання послуг. Власне, істотне зменшення часу доставки контенту і стало однією із найвагоміших переваг використання CDN мереж. Використання CDN також дає можливість суттєво економити трафік. Передати на інший континент файли один раз, зберігати їх там на локальному сервері й роздавати через локальні канали дешевше, аніж передавати той самий трафік тисячі разів через міжконтинентальні канали передавання даних. Всі ці методи необхідні для забезпечення ефективного використання ресурсів мережі та вузлів доставки контенту.

Для того, щоб швидко віддавати контент з локального сервера (Edge server) без звернення до оригінального сервера (Origin server), треба, щоб цей контент на ньому з'явився (і залишався). Існує багато схем кешування, найпоширеніші серед яких такі:

1. *Реплікація всього контенту.* Перевагою такого методу є велика швидкодія, навіть першої відповіді. До недоліків можна зарахувати високу вартість такої реалізації.
2. *Реплікація за першим зверненням (найпоширеніша схема).* Перше звернення повільне, а всі наступні швидші.
3. *Асинхронна реплікація за перевищенням певного порога звернень.* Економічніша версія, однак обслуговування більшості клієнтів повільне.

1.1.1 Основні терміни.

Головні терміни, які дають можливість розуміння CDN:

- 1) *CDN (Content Delivery Network)* - це географічно розподілена мережева інфраструктура, що забезпечує швидку доставку контенту користувачам веб-сервісів і сайтів. Що входять до складу CDN сервер географічно розташовуються таким чином, щоб зробити час відповіді для користувачів сайту сервісу мінімальним.

2) *Ориджін (origin)* - сервер, на якому зберігаються вихідні файли або дані, що передаються через CDN.

3) *PoP (point of presence, точка присутності)* - кешуючий сервер в складі CDN, розташований в певній географічній локації. Для позначення таких серверів також використовується термін *edge*.

4) *Динамічний контент* - контент, згенерований на сервері в момент отримання запиту (або змінений користувачем, або завантажується з бази даних).

5) *Статичний контент* - контент, що зберігається на сервері в незмінному вигляді (наприклад, бінарні файли, аудіо- та відеофайли, JS і CSS).

1.1.2 Призначення та сфера використання CDN.

Швидкий розвиток інтернет-технологій привів до ситуації, в якій звичайні сервери далеко не завжди можуть витримувати навантаження та якісно надавати сервіси кінцевим користувачам. Раніше був доволі поширеним термін “серверні ферми”, що являли собою сукупність серверів різних типів з однаковим географічним розташуванням. Тепер ці поняття відходять на задній план, а на зміну їм приходять хмарні (cloud) рішення та CDN мережі. Головне завдання, яке ставлять перед сучасними провайдерами послуг та контенту, – можливість масштабування мережевої інфраструктури в умовах постійного зростання трафіку та розподіл навантаження по різних локаціях. Для вирішення таких завдань використовують різні технології. Одна з них – створення мережі CDN, яка дає змогу зменшити навантаження, посилити захист від DDoS і прискорити завантаження сайту у віддалених локаціях.

Найчастіше CDN використовується для зменшення часу відгуку кешованого контенту, що, як ми вже згадували вище, зменшує можливість втрати відвідувачів через повільне завантаження ресурсу і тим самим скорочує можливі фінансові втрати. Також CDN допомагає знизити ризик

втрати доступу до контенту через падіння основного сервера. Контент буде доступний весь час, поки ви відновлює працездатність основного сервера.

Стрімінговий медіа-контент, розробники програмного забезпечення, сервіси доставки реклами, тематичні сайти є головними сферами використання CDN.

Ще один цікавий сценарій використання CDN - так званий live-streaming. Live-streaming: користувачі Інтернету з усього світу можуть в браузері (а іноді і в спеціальному додатку) дивитися або слухати трансляцію з місць подій.

Влаштовано це так: один або кілька оріджін-серверів приймають відеотрансляцію, яка відразу ж ретранслюється на точки присутності.

Оріджін-сервери при цьому контент клієнтам не роздають.

До складу стрімінгових CDN входять також балансувальники навантаження, що перенаправляють запити до найменш завантаженого на поточний момент edge-сервера.

1.2. Алгоритми розподілу запитів

Існують різні алгоритми, які можуть бути[6] використані для вибору сервера при маршрутизації запиту. Найбільш простими в реалізації є випадковий вибір (Random Choice - RC) та послідовний вибір (Round Robin - RR). Крім цих двох алгоритмів існують алгоритми, які вибирають найбільш відповідний сервер виходячи з деяких параметрів, таких як доступна пропускна здатність каналу (Bandwidth availability - BW), доступність носія інформації (Hard disk availability - HD) і кількість поточних підключень до сервера або доступність з'єднання (англ. connection availability - CON).

Коли система направляє запит користувача до одного з серверів з використанням одного з п'яти наведених алгоритмів, то пропускна здатність каналу від клієнта до сервера, завантаженість носія інформації і кількість підключень до сервера є трьома важливими факторами, за якими визначається, чи може сервер обробити запит клієнта. Якщо хоча б по

одному з цих параметрів сервер перевантажений, запит буде відхилено, а це означає, що CDN не зможе обробити запит клієнта. В архітектурі мереж доставки контенту мінімізація відсотка відхилених запитів найбільш важливий фактор при порівнянні продуктивності різних алгоритмів.

Алгоритм прийняття рішень з використанням нечіткої логіки:

Для даного алгоритму BW, HD і CON є вхідними параметрами. Покладаючись на ці три параметри алгоритм використовує механізм нечіткої логіки для прийняття рішення про вибір сервера. На першому етапі вхідні параметри перетворюються в відповідні значення нечіткої логіки відповідно до функцій приналежності. Визначаються функції приналежності для кожного вхідного параметра. Для параметра доступності каналу (BW) ця функція буде виглядати наступним чином. HBW - високе значення приналежності для пропускної здатності каналу, MBW - середнє значення і LBW - низьке значення відповідно.

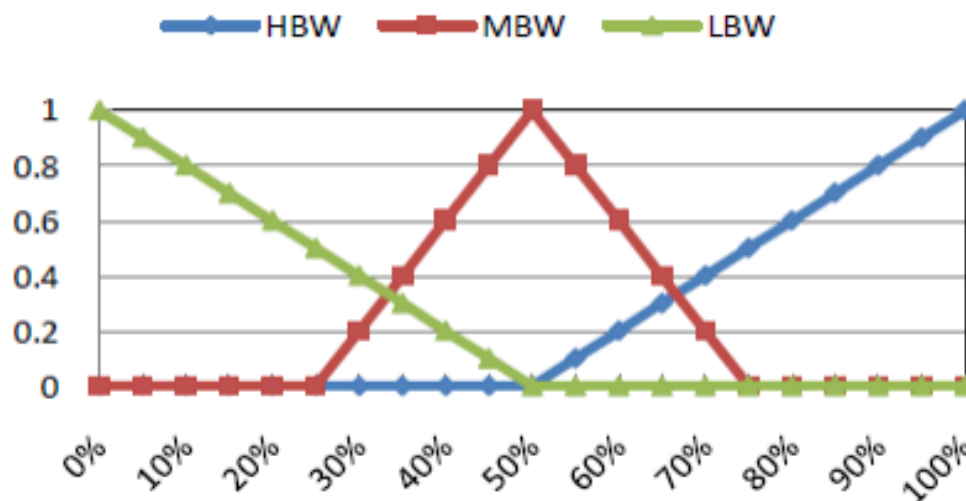


Рис. 1.2. Значення приналежності HND, MHD, LHD

На другому етапі відбувається обчислення правил. Є 9 значень приналежності (HBW, MBW, LBW, HND, MHD, LHD, HCON, MCON,

LCON).

Існує 27 можливих комбінацій цих значень. Для кожної з цих комбінацій ми визначаємо нечітке рішення з чотирьох можливих:

1. Настійно рекомендується сервіс - Так (Yes, Y)
2. Рекомендований сервіс - Можливо так (Probably yes, PY)
3. Не рекомендований сервіс - Можливо ні (Probably no, PN)
4. Настійно не рекомендується сервіс - Ні (No, N)

Кожне значення нечіткого рішення FD є мінімумом з трьох значень приналежності.

$$FD = \min \{HBW/MBW/LBW, HHD/MHD/LHD, HCON/MCON/LCON\} \quad (1.1)$$

Таблиця 1.1.
Прийняття рішення про вибір серверу

BW	HD	CON	FD
HBW	HHD	HCON	Y
HBW	HHD	MCON	Y
HBW	HHD	LCON	PY
HBW	MHD	HCON	Y
HBW	MHD	MCON	Y
HBW	MHD	LCON	PY
HBW	LHD	HCON	PY
HBW	LHD	MCON	PY
HBW	LHD	LCON	PN
MBW	HHD	HCON	PY
MBW	HHD	MCON	PY
MBW	HHD	LCON	PY
MBW	MHD	HCON	PY

Продовження таблиці 1.1.

MBW	MHD	MCON	PN
-----	-----	------	----

MBW	MHD	LCON	PN
MBW	LHD	HCON	PY
BW	HD	CON	FD
MBW	LHD	MCON	PN
MBW	LHD	LCON	PN
LBW	HHD	HCON	PY
LBW	HHD	MCON	PN
LBW	HHD	LCON	PN
LBW	MHD	HCON	PN
LBW	MHD	MCON	N
LBW	MHD	LCON	N
LBW	LHD	HCON	PN
LBW	LHD	MCON	N
LBW	LHD	LCON	N

Таким чином ми отримуємо 27 значень, котрі належать до чотирьох груп. З кожної групи береться мінімум. В результаті виходить чотири значення FD (Y), FD (PY), FD (PN) і FD (N).

На третьому етапі алгоритму, чотирьом значенням рішення призначається набір зважених значень. Кожне значення представляє різний набір значень. Підсумкове, "чітке" значення (Crisp Value - CV) обчислюється на основі зважених значень і значень нечітких рішень.

$$CV = \frac{\sum_i FD(i) * w(i)}{\sum_i w(i)} \quad (1.2)$$

Кожен сервер обчислює власне CV. Система розподілу запитів може використовувати CV для визначення найбільш відповідного сервера для обробки запиту. Сервер з найвищим CV є кращим для досягнення оптимальної продуктивності CDN.

Симуляція проводилася в порівнянні з алгоритмами RC, RR, BW, HD,

CON. По вертикальній осі графіка кількість відхилених запитів в відсотках. По горизонтальній осі середній час між запитами в мілісекундах.

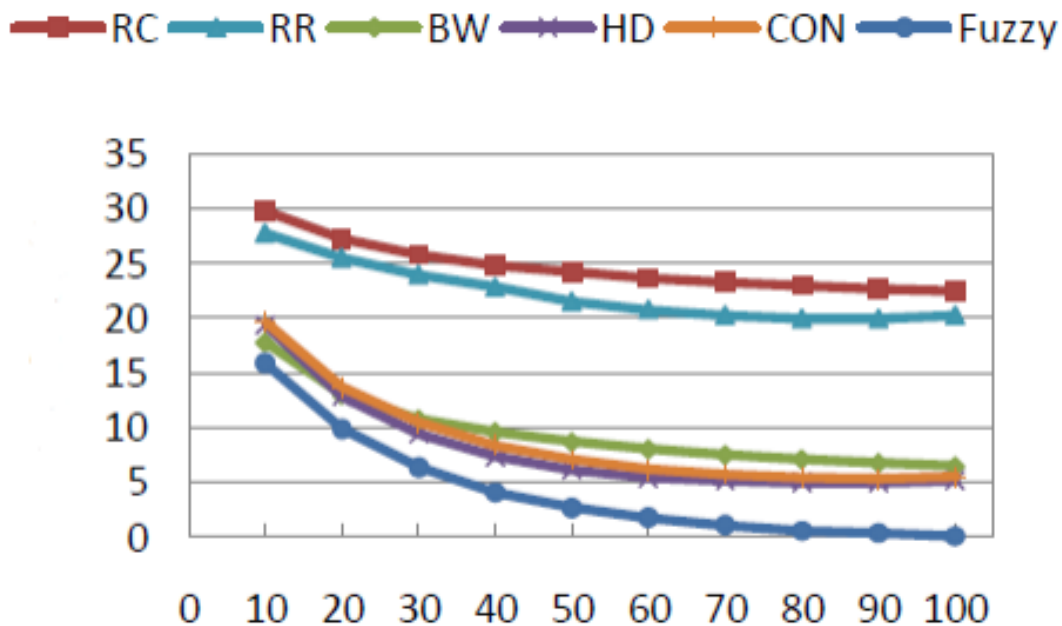


Рис 1.3. Порівняння алгоритмів RC, RR, BW, HD, CON

Оскільки, алгоритми RC і RR не використовують ніяких параметрів для розподілу запитів, для них відсоток відхилених запитів високий. При використанні тільки одного параметра (BW, HD, CON) відсоток відхилених запитів нижче ніж у RC і RR. Описаний алгоритм показує кращі результати, так як враховує відразу всі три параметра.

1.3. Принцип роботи CDN.

Принцип роботи CDN

Уявімо собі веб-сервіс, яким користуються люди по всій території України, а основні сервери розташовані в США, користувачі в цей час перебувають в самих різних географічних точках: скажімо, в Івано-Франківську (9 868 від Санта - Клара США), Одеса (10 282 км від Каліфорнії США), або Маріуполь (7 790 км від Нью-Йорка США). Чим далі користувач знаходиться від оригінального сервера, тим більше час «оригінальної» відповіді.

При використанні CDN все відбувається по-іншому: користувач з Івано-Франківська переадресовується до географічно найближчої кешуючої точки присутності в складі CDN, завдяки чому доставка статичного контенту відбувається набагато швидше.

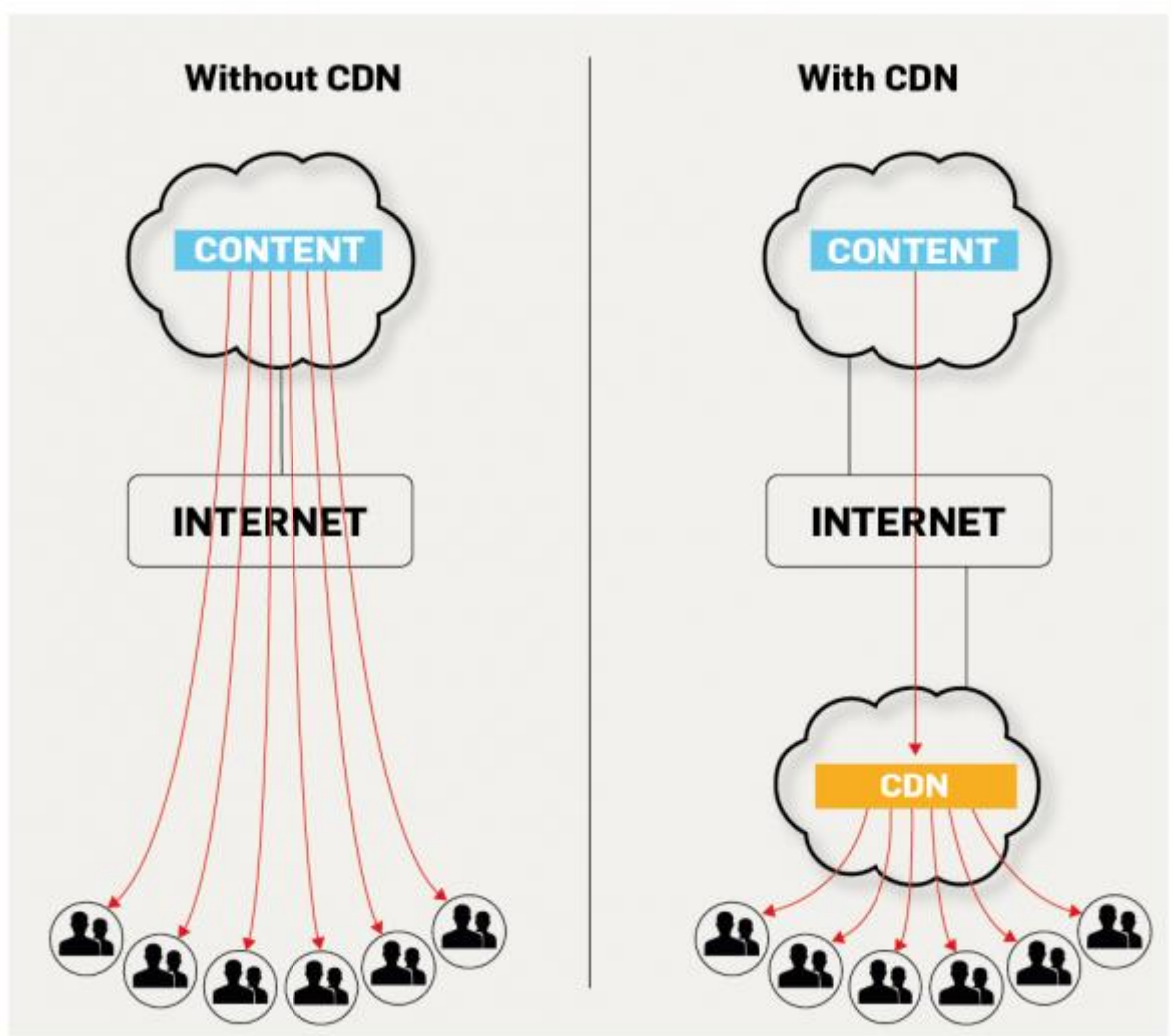


Рис. 1.4. Порівняльна схема доставки контенту.

Для прискорення роздачі динамічного контенту при використанні CDN, використовуються інші механізми: CDN-провайдер за рахунок своєї мережі скорочує мережевий маршрут.

Як організована передача даних?

Як правило, для настройки роздачі статичного контенту через CDN необхідно виконати наступні кроки:

1. Винести статику сайту на окремий домен, наприклад, `static.example.com` - це буде origin.

2. Для роботи через CDN створити домен виду `cdn.example.com`.
3. Підключити CDN у провайдера. Для підключення власнику веб-сервісу необхідно повідомити провайдера наступне:
 - а) домен, з якого він буде забирати статистику - `static.example.com`;
 - б) домен, з якого буде йти роздача - `cdn.example.com`.
4. У свого DNS-реєстратора налаштувати CNAME запис з `cdn.example.com` на домен CDN-провайдера, який CDN провайдер виділяє при підключенні. Наприклад, в CDN Selectel такий домен має вигляд `85e77c09-bc03-43bf-b8f3-9492ae33390f.selcdn.net`, де `85e72c09-bc03-43bf-b8f3-9492ae33390f` генерується автоматично.
5. На своєму сайті змінити домен для статистики, яку планується роздавати через CDN, на `cdn.example.com`.

Користувач набирає в рядку браузера адресу `www.example.com`, з якого він отримує HTML-сторінку. При цьому весь статичний контент, наприклад, графічні зображення, подгружається з CDN (з адреси `cdn.example.com`).

Веб-сервіс після підключення CDN буде працювати на тому ж оригінальному сервері. Кешовані частини сайту будуть завантажені на сервери CDN-мережі. Система шукає для користувача найближчий сервер та максимально швидко завантажує статистику сайту з нього.

Звернемо увагу на один важливий момент: сервери, що входять до складу CDN, не є подобою файлових серверів, на які контент розміщується для подальшого завантаження. CDN використовуються не для зберігання контенту, а для кешування на основі конкретних алгоритмів.

Як CDN визначає, де знаходиться PoP?

Як правило, для підвантаження контенту з CDN використовуються дві популярні технології: GeoDNS і AnyCast.

За допомогою GeoDNS можна прив'язати до одного доменного імені кілька IP-адрес. Залежно від географічного положення (визначається за IP-адресою, з якого прийшов запит) користувач перенаправляється на

найближчий сервер(PoP).

При використанні технології Anycast адреси загальні, але маршрутизація відбувається на «свої» сервери в межах регіону. При зверненні до адреси `www.example.com` користувач переадресовується на найближчу PoP. Провайдер користувача отримує кілька анонсів від різних мереж, в яких є точка присутності, і маршрутизатор провайдера вибирає з них найближчий. Відповідь аналогічним чином повертається по найбільш короткому маршруту.

Як кешуються дані в PoP?

Найпоширенішою є схема за першим зверненням: максимальну кількість часу на завантаження витрачає користувач, який звернувся до оригінального сервера першим. Усі наступні користувачі будуть отримувати дані, кешовані на найближчій до них точці присутності. Мозком CDN є система маршрутизації запитів. Завданням цієї системи входить вибір найбільш оптимального граничного вузла для обслуговування запиту клієнта. Вибір здійснюється на основі декількох параметрів, включаючи «близькість» вузла до клієнта, завантаження вузла, наявність в кеші необхідного контенту, а також інших можливих правил (наприклад, пов'язаних із захистом авторських прав або доступності послуг в певних географічних областях).

Тут дуже важлива географія: наприклад, після звернення користувача з Києва, дані будуть закешовані на сервері, що знаходиться на території України, що не вирішить проблеми зі швидкістю доступу для користувачів з Нью-Йорка або Каліфорнії.

Для подолання обмежень, що накладаються цією схемою, використовуються технології регіонального вилучення: сусідні сервери, що входять до складу CDN, забирають контент один в одного, а не звертаються до оригінального сервера.

У більшості випадках, користувач, що відправив запит на отримання статичного контенту, переадресовується до найближчої точки присутності і отримує кешовану версію цього контенту з неї. Якщо найближча точка

присутності не зможе знайти файли, почнеться пошук по сусіднім точкам присутності, звідки і буде перенаправлена відповідь користувачеві.

Як відбувається оптимальний розподіл контенту всередині CDN?

Змоделюємо топологію мережі у вигляді зваженого[6] неспрямованого графа $G = \{V, E\}$. Безліч вершин - безліч вузлів мережі, а кожна дуга в множині E , являє фізичне з'єднання між вузлами і зважено згідно деякої метриці, наприклад кількість хопів від одного вузла до іншого або інша більш складна функція, яка бере до уваги пропускну здатність каналу або завантаженість вузла. Визначимо дві підмножини V_a та V_R множини вузлів мережі. V_a - множина вузлів, на які потрапляють запити від користувачів. V_R - множина вузлів мережі, що містять контент.

Нехай існує C наборів даних. Користувачі можуть надсилати запит на доступ до даних одного з C наборів через вузли доступу з V_a , дані з цих наборів можуть розташовуватися на кожному з серверів V_R . Кожен вузол доступу може обслужити не більше ніж V_A^{MAX} запитів. Запити до заданого контенту обслуговуються вузлом з даними, що найбільше підходить. Віддача контенту через вузол на відстані більшій ніж поріг d_{max} вважається незадовільною. Кожен вузол з даними може обробити не більше K запитів. Не більше ніж V_R^{MAX} копій контенту може бути на одному вузлі даних.

Конфігурація запитів та вузлів з даними описується вектором стану x розміру $C(|V_a| + |V_R|)$:

$$x = (a, r) = (a_1^1, a_2^1, \dots, a_{|V_a|}^1, a_1^2, \dots, a_{|V_a|}^C, r_1^1, r_2^1, r_{|V_R|}^C) \quad (1.3)$$

кожна з змінних a_i^c відображає кількість запитів до контенту $c \in \{1, \dots, C\}$ на вузлі $i \in V_a$ та r_j^c кількість копій контенту $c \in \{1, \dots, C\}$ розташованих на вузлі $j \in V_R$.

Якість сервісу оцінюється функцією $A(x)$, яка визначається як сума відстаней між вузлом доступу і вузлом з даними по всім запитам

користувачів. Функція $M(x)$ оцінює вартість утримання копій та підтримки їх в актуальному стані та визначається, як:

$$M(x) = C_{Maint} \sum_{j \in V_R, c=1, \dots, C} r_j^c \quad (1.4)$$

1.4. Аналіз постачальників мереж доставки контенту і готових рішень для їх створення.

1. Akamai CDN

The Akamai Network - одна з найбільших в світі[6] розподілених обчислювальних платформ. Це мережа з більш ніж 105,000 серверів із спеціальним програмним забезпеченням, розташованих в 78 країнах. Сильною стороною мережі Akamai є алгоритми розподілу навантаження всередині мережі. Сервери знаходяться приблизно в 1,900 різних мережах, здійснюючи моніторинг Інтернету в реальному часі - збираючи інформацію про трафік, перевантаження, і проблемних точках. Akamai використовує зібрану інформацію для оптимізації маршрутів і динамічної реплікації, щоб доставляти контент швидше і надійніше. Основні способи оптимізації доставки контенту:

- 1) Мінімізація довжини маршрутів, за допомогою реплікації та доставки контенту з серверів, що знаходяться максимально близько до користувача.
- 2) Оптимізація маршрутів шляхом побудови шляхів через Інтернет, що обходять проблемні місця, стиснення контенту та реплікації пакетів.
- 3) Переміщення обчислень ближче до кінцевих користувачів, щоб уникнути затримок. (EdgeComputing).

2. Amazon CloudFront

Amazon CloudFront - веб-сервіс для доставки контенту. Amazon CloudFront інтегрується з іншими Amazon Web Service. Мета сервісу - дати розробникам і підприємствам простий спосіб поширювати контент для кінцевих користувачів з мінімальними затримками та високою швидкістю

передачі даних. Сервіс не є вільним для користування і входить в інфраструктуру сервісів Amazon Web Services.

3. CacheBoy

Cacheboy - надає відкриту платформу для здійснення ефективної доставки контенту. Дзеркальна інфраструктура відкритого проекту підтримується сервіс-провайдерам по всьому світу. Проект перенаправляє користувача на найближче дзеркало.

4. OpenCDN

OpenCDN - CDN рівня додатку, який створює репліку контенту і розділяє трансляції та записаний контент. OpenCDN написаний на Perl і використовує технологію Relay для поділу медіа пакетів. Зв'язок між вузлами і джерелами здійснюється під засобом XML - RPC.

5. P2P-Next

P2P-Next - програмне забезпечення наступного покоління для доставки контенту, яке може бути використане для одночасної передачі мільйонам людей. Воно використовує multicast для передачі даних мільйонам користувачів одночасно. Потік даних розподіляється по локальним серверам і прямий ефір може бути ретрансльований локальним глядачам. Так як IP роутери не підтримують multicast, оптимізовані механізми uni-cast, broadcast та multicast були адаптовані під P2P.

6. CloudFlare Railgun

Один із шляхів зниження навантаження всередині CDN - кешувати максимальна кількість контенту. Але в реальності тримати в кеші можна тільки близько 66% об'єктів, а решта 34% доводиться заново запитувати з сервера у разі оновлення. Для стиснення цього трафіку і був створений Railgun. В Railgun використовуються приблизно такі ж алгоритми, як при обробці послідовності кадрів у відео високої чіткості. Висока ступінь стиснення в відеокодеках досягається за рахунок стиснення не окремих

кадрів, а відмінностей між сусідніми кадрами. Це дозволяє стиснути кадр розміром в мільйони пікселів до майже 10 кілобайт. Теоретично його можна стиснути взагалі в один байт, якщо він нічим не відрізняється від попереднього кадру. При зміні веб-сторінки, змінюється тільки невелика частина вмісту, і досить передати зміни між актуальною версією і тієї, яку клієнт отримав в попередній раз. Технічно, Railgun складається з двох компонентів: відправника (sender) і одержувача (listener). Відправники встановлені в кожному дата-центрі CDN-мережі. Одержувач — програмний компонент, який встановлюється на прикордонні вузли мережі. Між відправником і одержувачем встановлюється TCP-з'єднання, захищене TLS, за яким бінарний протокол Railgun здійснює асинхронну передачу HTTP-запитів. Для веб-клієнта система Railgun виглядає як проксі-сервер, хоча насправді це спеціалізована система зі специфічними функціями. Одна з них — стиснення контенту, який неможливо кешувати, за рахунок синхронізації версій сторінок. При оновленні версії сторінки по мережі передається тільки зміна між попередньою і новою версією. Дослідження показали, що максимальне стиснення досягається на новинних сайтах з великою відвідуваністю. Наприклад, бінарне порівняння головної сторінки сайту reddit.com показує зміну в середньому 2,15% протягом 5 хвилин і 3,16% протягом години. Головна сторінка The New York Times бінарно змінюється на 0,6% за п'ять хвилин і на 3% протягом години. Для головної сторінки BBC News ці показники становлять 0,4% і 2%, відповідно. У загальному випадку, для найбільш популярних сайтів при повторному запиті змін різниця займає один пакет TCP.

Висновки:

1. Сформовано принцип доставки контенту та процес отримання контенту за допомогою CDN.
2. Розкрито основні функції та сфери застосування сервісу швидкої доставки контенту.

3. Визначено алгоритми кешування, які є фундаментом для отримання контенту з локального сервера без звернення до оригінального сервера.

4. Проаналізовано постачальників CDN і готових рішень для їх створення.

РОЗДІЛ 2.

ОГЛЯД ТЕХНОЛОГІЙ ПЕРЕФЕРІЙНОГО ОБЧИСЛЕННЯ. EDGE COMPUTING

Хмарні обчислення вже кілька років, як стали буденністю в системах розвинених країн. Найгарячіша тема тепер - периферійні (або прикордонні, або крайові) обчислення, edge computing.

Глобальний обсяг даних за минулий рік склав 33 зеттабайт, за даними IDC, при цьому висхідний тренд триває, і до 2025 року буде генерувати 175 зеттабайт. Ці дані повинні бути оброблені миттєво. Після обробки, частина даних буде знищена по причині недостатньо пам'яті на збереження, а частина збережена для подальшого ретроспективного аналізу та інше. В теорії все легко, але практично, на даний час, наявна інфраструктура потребує вдосконалення., тому вимагає як кількісного розвитку, так і якісного.

Граничні обчислення (англ. Edge computing, периферійні обчислення) - це парадигма розподілених обчислень, що здійснюються в межах досяжності кінцевих пристроїв. Даний тип обчислень використовується для скорочення часу мережевого відгуку, а також більш ефективного використання пропускної здатності мережі.

Периферійні обчислення - вони ж Edge computing, являють собою своєрідний «кеш» для даних і, звичайно ж, для обчислень, причому в найширшому розумінні.

2.1. Призначення технологій периферійних обчислень.

Метою Edge computing є перенесення обробного додатку або точки забезпечення загальної функціональності ближче до джерела дії, туди, де технологія розподілених систем взаємодіє з фізичним світом. Методологія Edge (Fog) Computing це забезпечення близькості до кінцевих пристроїв, географічний розподіл обчислювальних потужностей і підтримка мобільності. Такий підхід дозволяє знизити навантаження на мережу і ЦОД, зменшує затримки в обслуговуванні і покращує якість сервісів. Засоби Edge computing не потребують контакту з будь-якою централізованою хмарою,

хоча і може взаємодіяти з ним при необхідності. Таким чином, хмарні обчислення працюють на «великих даних», в централізованих точках консолідації інформації, в той час як Edge Computing працює на «миттєвих даних», тобто на поточних даних, що генеруються датчиками або контрольованими системами.

Граничні обчислення в розрізі управління вхідними IoT-даними вплинуть практично на всі компанії майже в кожній галузі економіки та державного сектору. Edge computing охопить всі сфери діяльності, починаючи з автоматизації пропускнуго контролю, збору даних про якість вироблених товарів, моніторингу руху транспортних засобів і закінчуючи роботизацією заводів і дистанційною хірургією.

Непрямою вказівкою про наближення IoT-ери є модулі та додаткові можливості для збору або аналізу IoT-даних, які вбудовують в свій софт постачальники ПО. Втім, це не означає, що підприємствам слід негайно купувати дорогі ПО і IoT-мережі, багато хто з цих ресурсів пропонує за цінами нижче, ніж провайдери хмарного сервісу.

2.2 Застосування технологій периферійних обчислень у мобільних мережах

Відкриття цього середовища IT-послуг, дозволить додаткам і сервісам від мобільних операторів, постачальників послуг і контенту ефективно і без проблем інтегруватися в багатопроцесорні мобільні платформи для мобільних обчислень. Характеристики та можливості, пропоновані платформою MEC, можуть бути використані, щоб забезпечити близькість, контекст, гнучкість і швидкість для більш широких інновацій, які можуть бути переведені в унікальну цінність і дохід.

Доступ до контенту і додатків може бути прискорений; їх чуйність може бути збільшена, максимізуючи швидкість та інтерактивність. Популярний і локально релевантний контент може доставлятися безпосередньо, коли користувачі підключаються, обмежуючи смугу пропускання до ядра і хмари.

Знання умов радіозв'язку в режимі реального часу і контекстної інформації може використовуватися для оптимізації роботи мережі і обслуговування (реагування і адаптації до умов, що змінюються мережі). Це поліпшить якість обслуговування та використання мережневих ресурсів, що дозволить їм ефективно обробляти збільшені обсяги трафіку. Мережа реального часу і маломасштабна контекстна інформація (включаючи місцезнаходження) можуть використовуватися для збагачення мобільного широкосмугового доступу шляхом створення персоналізованих сервісів, орієнтованих на індивідуальні потреби і переваги.

Граничні обчислення (MEC), раніше відомі як мобільні граничні обчислення, - це тип граничних обчислень, який розширює можливості хмарних обчислень, виводячи їх на кордон мережі та розраховані на велику кількість користувачів.

У той час як традиційні хмарні обчислення виконуються на віддалених серверах, розташованих далеко від користувача і пристрою, MEC дозволяє процесам проходити в базових станціях, центральних офісах і інших точках агрегації в мережі.

Зміщуючи навантаження хмарних обчислень на окремі локальні сервери, MEC допомагає знизити перевантаження в мобільних мережах і зменшити затримки, підвищуючи якість обслуговування (QoE) для кінцевих користувачів.

Варіантів застосування Edge computing, зауважимо, можливо досить багато, причому вже не теоретичних, а реалізованих на ринках споживачів.

Багато хто каже, що технологія **Edge computing** може стати заміною для всіх відомого “хмарного обчислення”.

Кращий спосіб продемонструвати використання методу Edge Computing - через кілька ключових прикладів. Ось кілька сценаріїв, де граничні обчислення найбільш корисні:

1) Автономні транспортні засоби

Автомобілі з автономним приводом або інтеграцією штучного

інтелекту та інші транспортні засоби, потребують величезний обсяг даних зі свого оточення для правильної роботи в реальному часі. Затримка відбудеться, якщо будуть використовуватися хмарні обчислення.

2) *Потокові сервіс*

Такі сервіси, як Netflix, Hulu, Amazon Prime і Disney +, створюють велике навантаження на мережеву інфраструктуру. Граничні обчислення допомагають створити більш оптимальний підхід за допомогою прикордонного кешування: популярний контент кеширується на об'єктах, розташованих ближче до кінцевих користувачів для більш швидкого доступу.

3) *Розумні будинки*

Подібно потоковим послугам, зростаюча популярність розумних будинків, створює певні проблеми. Існує дуже велике навантаження на мережу, що не дозволяє покладатися тільки на хмарні обчислення. Обробка інформації ближче до джерела означає меншу затримку і більш швидкий час відгуку в аварійних ситуаціях. Прикладами можуть служити реагування на аварійну подію, медичними бригадами, пожежниками або службою поліції.

2.2.1. Застосування у мережах 2G.

Фундаментом для розвитку 2G мережі, стало створення глобального стандарту цифрового мобільного стільникового зв'язку, GSM(Global System for Mobile Communications), за часом (TDMA) та частоті (FDMA). Таким чином, ми отримали можливість передачі SMS та мультимедійних повідомлень або MMS.

Технологія GSM(Global System for Mobile Communications) дала спосіб використання радіочастот, коли в одному частотному інтервалі знаходяться кілька абонентів, різні абоненти використовують різні тимчасові слоти (інтервали) для передачі.

Отримання контенту в мобільній мережі 2G, стало можливим завдяки технології WAP(*Wireless Application Protocol*).

Додаток Wireless Application Protocol (WAP) надає мобільним

абонентам можливість доступу в Інтернет і є універсальним, відкритим стандартом бездротових телефонів і PDA для доставки контенту Інтернет та інших додаткових послуг. Важливо відзначити, що хоча WAP дозволяє користувачеві передавати і приймати текст, цей додаток і є технологією 2.5G для високошвидкісної бездротової доставки даних.

Технології Edge Computing, як такої, не було в застосуванні 2G мережі.

2.2.2. Застосування у мережах 3G.

3G (third generation – третє покоління) – набір послуг, який поєднує як високошвидкісний мобільний доступ в мережу Інтернет, так і технологію радіозв'язку, яка створює канал передачі даних. 3G технологія прийшла на зміну мобільних мереж GSM і CDMA.

В мережах третього покоління, є важлива перевага – це покращений захист від обривів зв'язку під час руху користувача, за рахунок використання методу так званого м'якого «хэндовера» (англ. handover).

Суть методу полягає в тому, що по мірі віддалення від однієї базової станції, клієнта підхоплює інша. Після підключення клієнта, базова станція починає передавати все більше і більше інформації, у той час як перша станція передає все менше, поки клієнт взагалі не покине зону її обслуговування.

Мережі 3G в основному використовуються абонентами в якості засобу для підключення до глобальної мережі Інтернет мобільних телефонів з метою доступу до мультимедійних послуг зв'язку: голосові і відеодзвінки, завантаження і скачування даних, відвідування сайтів в мережі, онлайн-телебачення так контенту різного формату.

Головною особливістю і безперечною перевагою 3G є високошвидкісна передача даних. Основою мобільного зв'язку третього покоління є технологія IP, яка базується на пакетній передачі даних. Це означає, що абонент може безперервно перебувати в режимі онлайн.

Найбільш перспективними послугами на базі 3G вважалися

відеотелефонія, потокове відео, а також бізнес-послуги, в тому числі пошта та банківські сервіси.

Технологією доставки контенту в мережі 3G є MBMS - режим широкомовної передачі мультимедійних даних (Multimedia Broadcast Service).

MBMS дозволяє контролювати призначені для користувача дані і управляти ними. Залежно від числа користувачів виділяються два режими MBMS: двоточковий режим і широкомовний режим («точка-многоточка»).

У першому випадку передача мультимедійних даних (потокове відео і аудіо, завантаження файлів, повторні передачі на вимогу) здійснюється для одного користувача або для невеликої групи користувачів. У другому випадку здійснюється широкомовна трансляція цих даних. Швидкість передачі даних при цьому становить 64 кбіт / с.

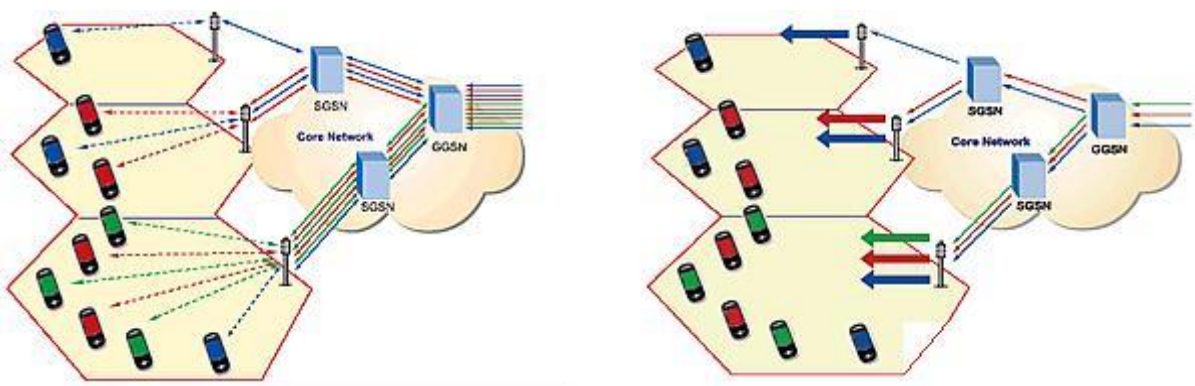


Рис. 2.1. Схема широкомовної передачі даних MBMS

Основною технологією доставки контенту в мережі 3G є режим широкомовної передачі мультимедійних даних. Технологія периферійних обчислень не могла бути застосована в мережі третього покоління, через відсутність необхідної потужності мережі та діапазонів частот.

2.2.3. Застосування у мережах 4G/LTE/4G+/LTE-A.

4G - покоління мобільного зв'язку з підвищеними вимогами. До четвертого покоління прийнято відносити перспективні технології, що дозволяють здійснювати передачу даних зі швидкістю, що перевищує 100 Мбіт / с - рухомим (з високою мобільністю) і 1 Гбіт / с - стаціонарним абонентам (з низькою мобільністю).

Основна відмінність 4G від 3G - велика швидкість передачі даних в мережах, необхідна для більш комфортного використання сучасних сервісів стрімінга музики, відео, онлайн ігор. Крім того, стандарти 4G повністю засновані на пакетній передачі даних, а ще такі мережі відрізняються більшою ємністю (одна базова станція може працювати з великим числом абонентів) і пропонують кращу стабільність з'єднання.

Повноцінне покоління 4G з'явилося з виходом так званого 4G + або LTE Advanced. Основні принципові нововведення, які відрізняють LTE-A від LTE - це агрегація частот (CA - Carrier Aggregation), поліпшене використання багатоантенних технологій (MIMO), а також підтримка релейного режиму включення базових станцій (RN - Relay Nodes).

Технологія LTE Broadcast, или eMBMS є основним методом доставки контенту в мережі LTE.

LTE Broadcast, або eMBMS, - це трансляція одних і тих самих даних практично необмеженій кількості користувачів одночасно. Трансляція здійснюється в обрану географічну зону з використанням одного частотного діапазону в один часовий період. Таким чином забезпечується оптимальне використання мережевого ресурсу і досягається висока якість демонстрації відео в мережах LTE.

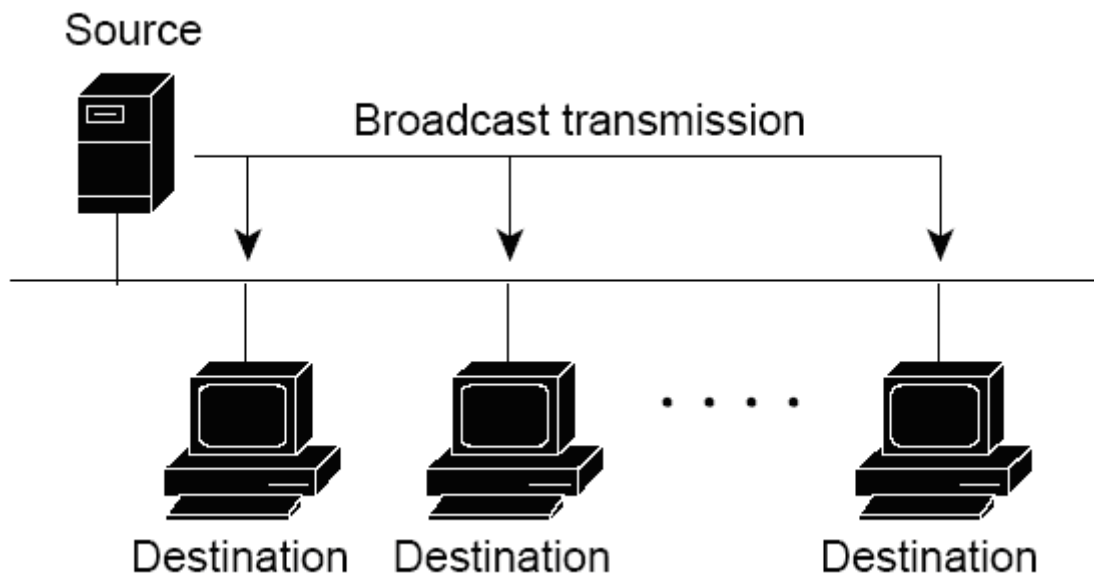


Рис 2.2. Приклад передачі трафіку на основі LTE Broadcast.

Для передачі пакетів даних застосовуються дві технології: Broadcast (широкомовне мовлення) і Unicast (індивідуальне мовлення). Обидві технології працюють в рамках однієї соти і на загальних потужностях, при цьому доступні ресурси мережі динамічно розподіляються між каналами Broadcast і Unicast. Одночасне використання Broadcast і Unicast розширює можливості для передачі відео в мережі LTE.

Оскільки, технологія Edge computing – це розробка, яка виникла під час еволюції мобільних станцій, телекомунікаційних мереж та використання її прийшло тільки в мережі п'ятого покоління.

2.2.4. Застосування у мережах 5G.

Стандарт мобільного зв'язку п'ятого покоління (5G) - це новий етап розвитку технологій, який покликаний розширити можливості доступу в Інтернет через мережі радіодоступу.

Атуальність хмарних обчислень для мобільних мереж знаходиться на підйомі. Сервіси соціальних мереж, такі як Facebook і Twitter, контент з YouTube і Netflix, а також інструменти навігації з Google Maps - все це в хмарах. Крім того, зростаюча залежність користувачів від мобільних пристроїв для виконання обчислювально-ресурсномістких операцій, як

особистих, так і пов'язаних з бізнесом, вимагає розвантаження хмари для досягнення кращої продуктивності, що продовжує термін служби батареї. Такі цілі було б складно і дорого реалізувати, не наближаючи хмарний сервер до кордону мережі і користувача. У відповідь на цю вимогу оператори мобільного зв'язку працюють над Mobile Edge Computing (MEC), в якій обчислювальні ресурси, сховища і мережеві ресурси інтегровані з базовою станцією. Обчислювальні додатки, чутливі до затримок і додатків, такі як доповнена реальність і обробка зображень, можуть бути розміщені на кордоні мережі.

MEC - це хмарне середовище IT-сервісів, в якій додаток працює ближче до клієнтів. Таким чином, час очікування і перевантаження мережі можуть бути зменшені, що зазвичай дає більш високу продуктивність.

Додаток MEC розміщується в розподіленій хмарі системи MEC. Він може належати одному або декільком сегментам мережі (тобто. Віртуальним мереж), які були налаштовані всередині базової мережі 5G.

Тим самим, ми бачимо переваги edge computing. Якщо немає справжньої необхідності збирати всі дані в централізованому хмарному сховищі, немає сенсу витратити кошовну пропускну здатність на їх транспортування.

Насправді, повністю ефективний дизайн IoT може бути лише один - датчики підключаються до хмари, коли у них є щось важливе для передачі. Такий підхід дозволяє скоротити витрати на IoT-мережі за рахунок використання технологій, які використовують більш дешевий метод білінгу з оплатою за кілобіт, а не більш дорожчий через постійне підключення.

При використанні граничних обчислень слід враховувати, що, оскільки дані не зберігаються довготривало, вони в кінцевому підсумку видаляються з периферійних пристроїв, і це не сприяє глибинному аналізу великих даних. Пам'ятайте, що прикордонні пристрої надають тільки результати обробки локально зібраних даних. Таким чином, якщо ваш IoT-проект вимагає, щоб ви зберігали всі зібрані первинні дані для прийняття рішення і аналітики, то

edge-обчислення не підходять.

Багато застосувань 5G покладаються[2] на крайові обчислення для взаємодії в режимі реального часу, локальної обробки, високої швидкості передачі даних, та висока доступність, включаючи:

1) Охорона здоров'я;

Наприклад віддалена хірургія та діагностика, а також моніторинг життєвих ознак пацієнта та дані. Лікарі можуть використовувати віддалену платформу для роботи хірургічними інструментами, щоб врятувати життя на відстані.

2) Розважальні та мультимедійні програми, такі як потокове HDTV або 3D TV.

3) Віртуальна реальність, доповнена реальність і змішана реальність;

Наприклад, потокове передавання відеовмісту до окуляр віртуальної реальності.

4) Тактильний Інтернет, який є наступною еволюцією Інтернету речей, забезпечує надто чутливий та надто надійний мережевий зв'язок для забезпечення успішної доставки повідомлень контролю в реальному часі та фізичних тактильних переживання віддалено.

5) URLLC, що забезпечує високу надійність між UE зокрема в комунікаціях M2M, підтримує низькі затримки передачі невеликих корисних навантажень з дуже високої надійністю від обмеженого набору UE, таких як пожежна сигналізація.

6) Інтернет речей, таких як розумні прилади, які під'єднуються до пристрою (наприклад, побутова техніка) використовуючи Інтернет.

7) Фабрики майбутнього, такі як розумні машини. Оператори можуть використовувати віддалену платформу для роботи важких машин, особливо тих, що розташовані у важкодоступних та небезпечних місця, з безпечного та зручного місця.

Основні ролі, які виконує Edge computing в мережі 5G:

1. *Місьцеве сховище.*

Крайові обчислення вивантажують[1] величезну кількість даних до крайових хмар. Прикладами даних, що зберігаються, є обчислювальні стратегії (наприклад, стратегія завантаження обчислень), метадані (наприклад, часові позначки та географічне розташування) та дані моніторингу. Крайовий сервер пропонує різні типи стратегій зберігання для підтримки різних типів даних. Наприклад, ефемерне зберігання забезпечує тимчасове зберігання даних для набору взаємопов'язаних мобільних пристроїв.

2. *Локальні обчислення.*

Крайові обчислення вивантажують обчислення та процес від менш складних (наприклад, смартфонів) та дуже складних (наприклад, хірургічних інструментів та розумних заводів) до крайових хмар, яка забезпечує можливості локального обчислення та обробки даних, близьких до UE, незалежно та автономно. Перевага полягає в тому, що крайові хмари виконують невеликі завдання та надають відповіді в режимі реального часу на місцях, що сприяє зниженню витрат та затримці, що виникають для надсилання потрібних даних у хмару.

3. *Локальний аналіз даних.*

Крайові обчислювальні процеси, що здійснюють критичний та реальний аналіз даних у величезній кількості необроблених даних, зібраних із різних програм у безпосередній близькості для отримання цінної інформації. Можливість робити аналіз даних локально зменшує затримку, необхідну для надсилання даних у хмару, а також чекати відповіді з хмари. Згодом результати аналізу місцевих даних використовуються для прийняття рішень.

4. *Локальне прийняття рішень.*

Крайові обчислення допомагають суб'єктам приймати рішення в реальному часі та відповідні дії в автоматизованому порядку на основі добре оброблених даних. Можливість приймати рішення на місцевому рівні

знижує залучення більшої кількості компонентів та обміну даними чи інформацією, що призводить до:

- а. поліпшення доступності системи, зокрема хмари;
- б. покращена пропускна здатність.

Наприклад, обчислювальна техніка полегшує місцеве прийняття рішень на автоматизованих заводах. Кілька організацій можуть приймати рішення спільно.

5. *Локальні операції.*

Обчислення по краях дозволяє віддалено керувати та контролювати - особливо критичні пристрої, включаючи пристрої, що знаходяться в небезпечному середовищі, - з відстані або більш зручного або безпечного місця.

б. *Поліпшення локальної безпеки.*

Крайові обчислення слугує додатковим шаром між хмарою та підключеними пристроями з метою покращення безпеки мережі, включаючи UE з обмеженими ресурсами. Крайові хмари можуть слугувати захищеними розподіленими платформами, що забезпечують управління обліковими даними безпеки, виявлення зловмисного програмного забезпечення, розподіл патчів програмного забезпечення та надійні комунікації для виявлення, перевірки та протидії атак. Перевага полягає в тому, що завдяки близькій близькості до обчислювальних технологій зловмисні об'єкти можуть швидко виявлятися та виділятися, і реалізовувати реакції в режимі реального часу для покращення наслідків атак. Це допомагає мінімізувати збої в обслуговуванні. Крім того, масштабованість та модульність, а також можливості крайових обчислень можуть сприяти розгортанню блокової ланцюга з обмеженими можливостями.

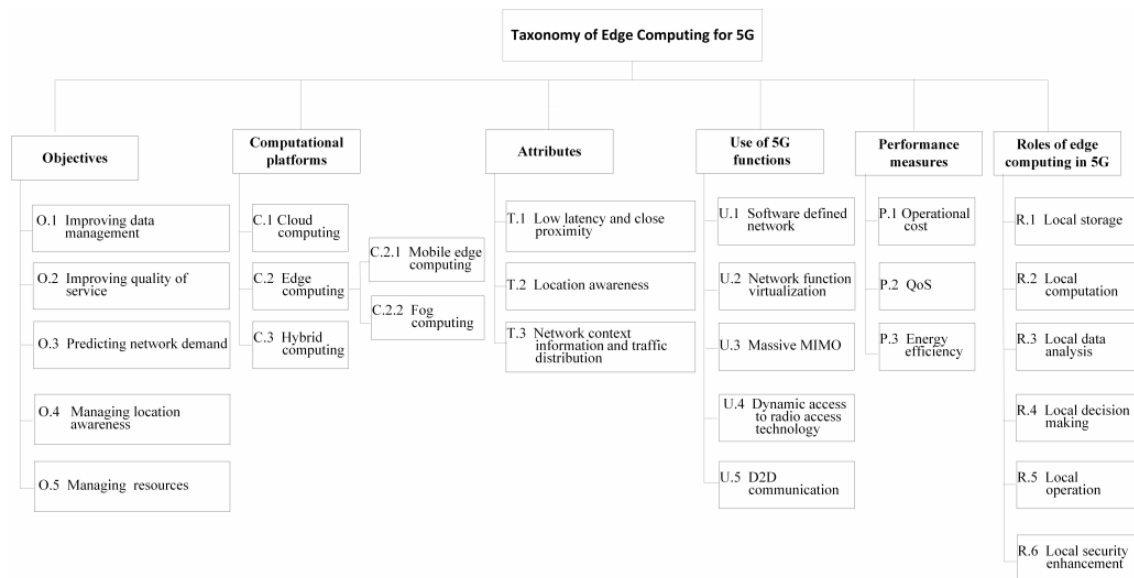


Рис. 2.3. Таксономія Edge Computing в 5G мережі

2.3 Огляд можливостей мережі 5G для побудови CDN для поширення контенту між мобільними пристроями.

CDN в 5G, дає операторам більше контролю[4] над тим, як вміст запиту маршрутизується через їхні мережі та враховує їх власні топології та ефективність. У цьому підході інфраструктура розподілу розміщується в мережі оператора, яка може бути або активною (вплив на чприбуток), або пасивною (без впливу на дохід) CDN.

Активний CDN дозволить операторам отримувати дохід від постачальників послуг контенту, розміщуючи необхідну інфраструктуру в їхній мережі.

Активний CDN означає, що необхідний мережевий елемент CDN може бути розміщений в RAN стільникового мобільного оператора або в ядрі пакета, як показано на рисунку 2.4.. Активний CDN може покращити QoS кінцевих користувачів та зменшити інвестиційні витрати в мобільну мережу, а також створити нові активні потоки доходу, зокрема в поєднанні з QoS мобільної мережі (контролер політики).

Пасивний CDN більше схожий на прозоре кешування, коли мережевий елемент CDN отримує трафік із зовнішніх джерел. Цей тип кешування допомагає зменшити транзитні витрати на IP-адреси та перингові витрати.

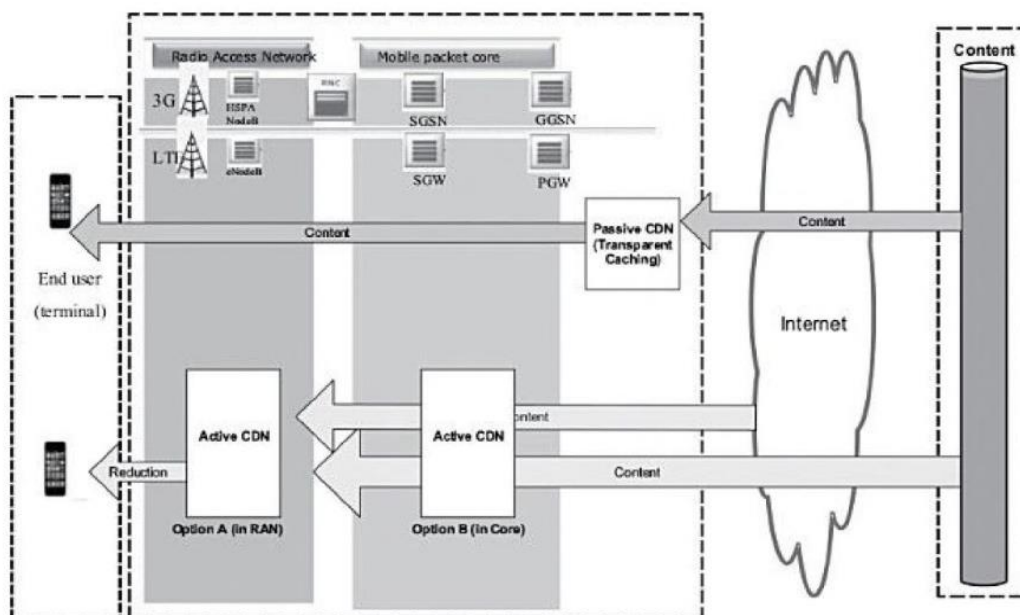


Рис. 2.4. CDN в мобільних мережах

CDN покращує продуктивність мереж, переміщуючи вміст ближче до кінцевого користувача. Деякі випадки, коли CDN може змінити ситуацію, є наступними:

1) Як механізм оптимізації додатків, CDN може оптимізувати кількість даних, що надсилаються, збільшуючи використання наявної пропускної здатності користувача. Більше того, коли CDN підключає користувача до географічно найближчого сервера, чистий ефект полягає у скороченні затримки, запровадженої мережею. Такі двигуни здійснюють оптимізацію на різних рівнях мережі для прискорення роботи доставки вмісту для всіх типів користувачів.

2) Як реплікатор файлів, особливо коли потрібні великі файли даних у кількох географіях, CDN можуть своєчасно копіювати дані з корпоративного штабу в одному регіоні до одного або декількох регіональних штаб-квартир.

3) Як найвище джерело пропускної здатності. CDN можна використовувати для пом'якшення ефекту високого обсягу, бурхливого та

нераціонального зростання трафіку та підтримки продуктивності програми на прийнятному рівні.

Висновки:

1. Проведено аналіз можливостей периферійних обчислень та визначені галузі застосування технології.
2. Проаналізовано можливості застосування технології Edge computing в мобільних мережах.
3. Розкрито головні поняття ролі граничних обчислень в мережі 5G.
4. Визначено можливості CDN та 5G для поширення контенту між мобільними пристроями.

ЗАГАЛЬНІ ВИСНОВКИ ПО РОБОТІ

1. Оглядовий аналіз технології CDN, принципів роботи мережі, сфер застосування, дозволив чітко визначити головні функції та задачі мережі доставки контенту. Основною задачею мобільних CDN є інтегрування послуги мобільної доставки, які оптимізують доставку будь-якого виду контенту.

2. З аналізу технології Edge computing та огляду сфер застосування у мобільних мережах різного рівня було визначено, що граничні обчислення розширюють можливості хмарних обчислень, виводячи їх на кордон мережі та розраховані на велику кількість користувачів.

3. Проведена оцінка можливостей 5G в мережі доставки контенту дозволяє робити висновок, що обчислювальні прикладні програми, чутливі до затримок і прикладні програми такі як доповнена реальність і обробка зображень можуть бути розміщені на кордоні мережі завдяки мережі п'ятого покоління.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Saad Asif., “5G Mobile Communications: Concepts and Technologies” – 2019
2. Najmul Hassan, KOK-LIM ALVIN YAU ,Celimuge Wu, “Edge Computing in 5G: A Review” – 2019
3. Samuilov K., Vasilevsky V., Korolkova A., Shalimov I., Vasin N., Kulyabov D., “Networks and telecommunications. Textbook and workshop for SPO” – 2019
4. Muthann A., Atey A., Filimonov M., “Study of cloud computing in cellular systems” – 2017
5. Smelyansky R., “Hierarchical Peripheral Computing” – 2019
6. Likhobabin S., “Content Delivery Network Load Balancing Algorithms” – 2013